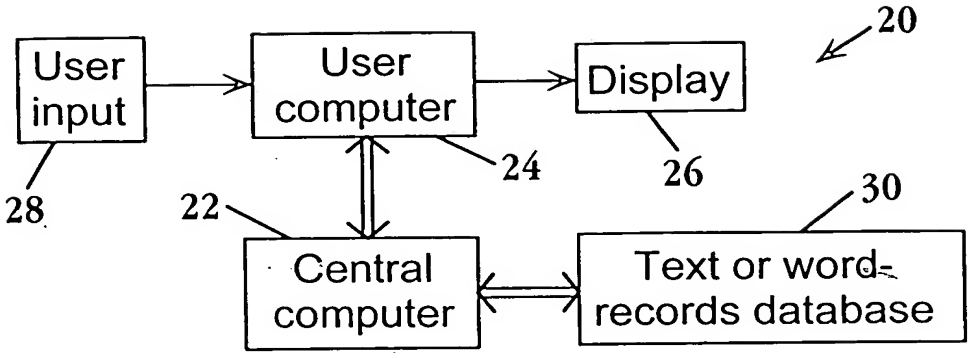

20
User input
User computer
Display
28
24
26
22
30
Central computer
Text or word-records database

Fig. 1

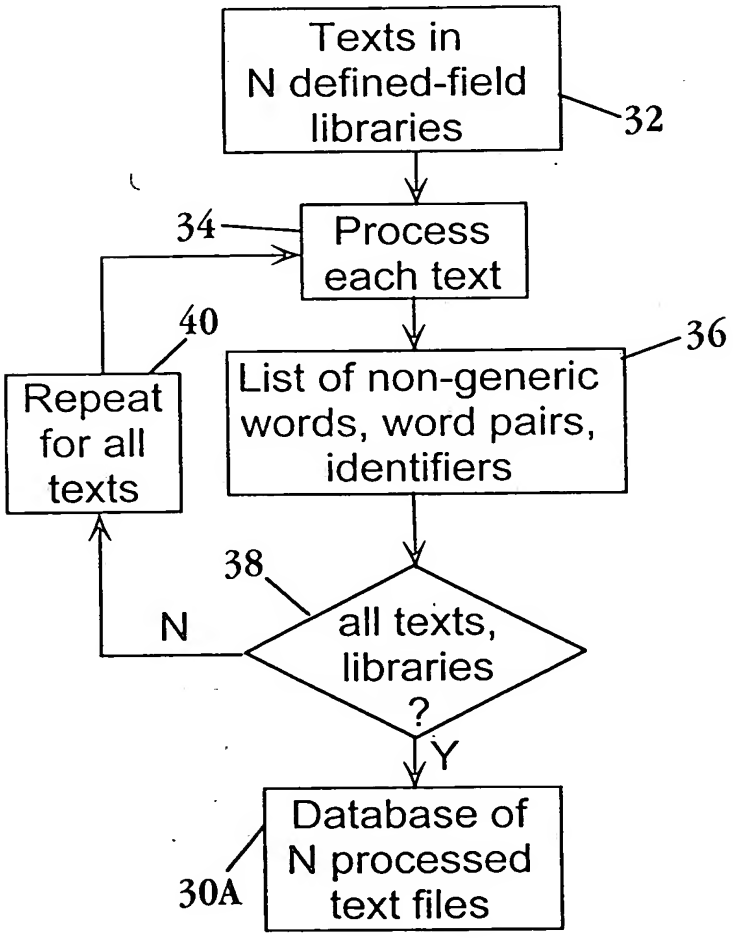

Texts in N defined-field libraries
32
34
Process each text
40
36
Repeat for all texts
List of non-generic words, word pairs, identifiers
38
N
all texts, libraries ?
Y
Database of N processed text files
30A

Fig. 2A

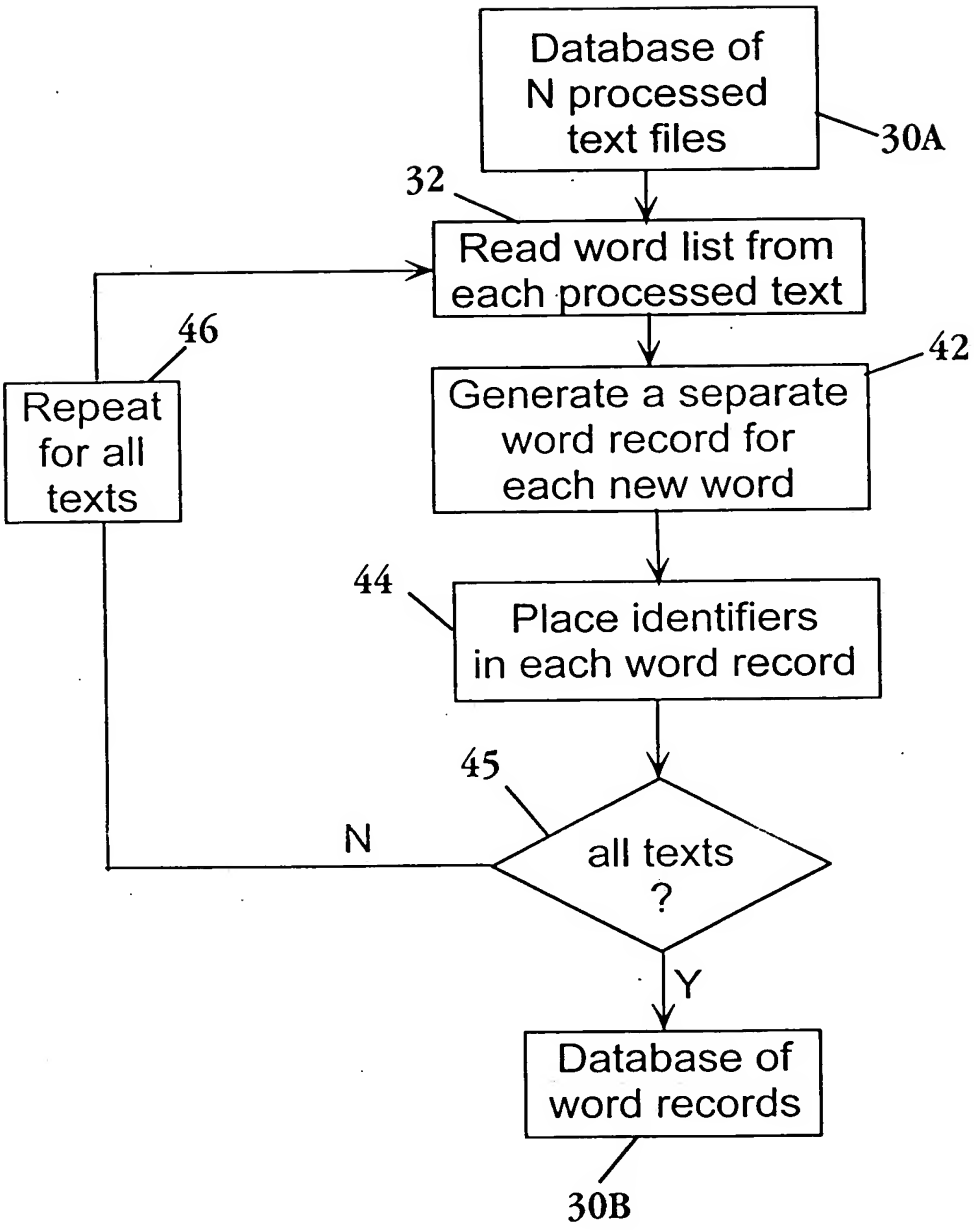

Database of N processed text files
30A
32
Read word list from each processed text
46
Repeat for all texts
42
Generate a separate word record for each new word
44
Place identifiers in each word record
45
N
all texts ?
Y
Database of word records
30B

Fig. 2B

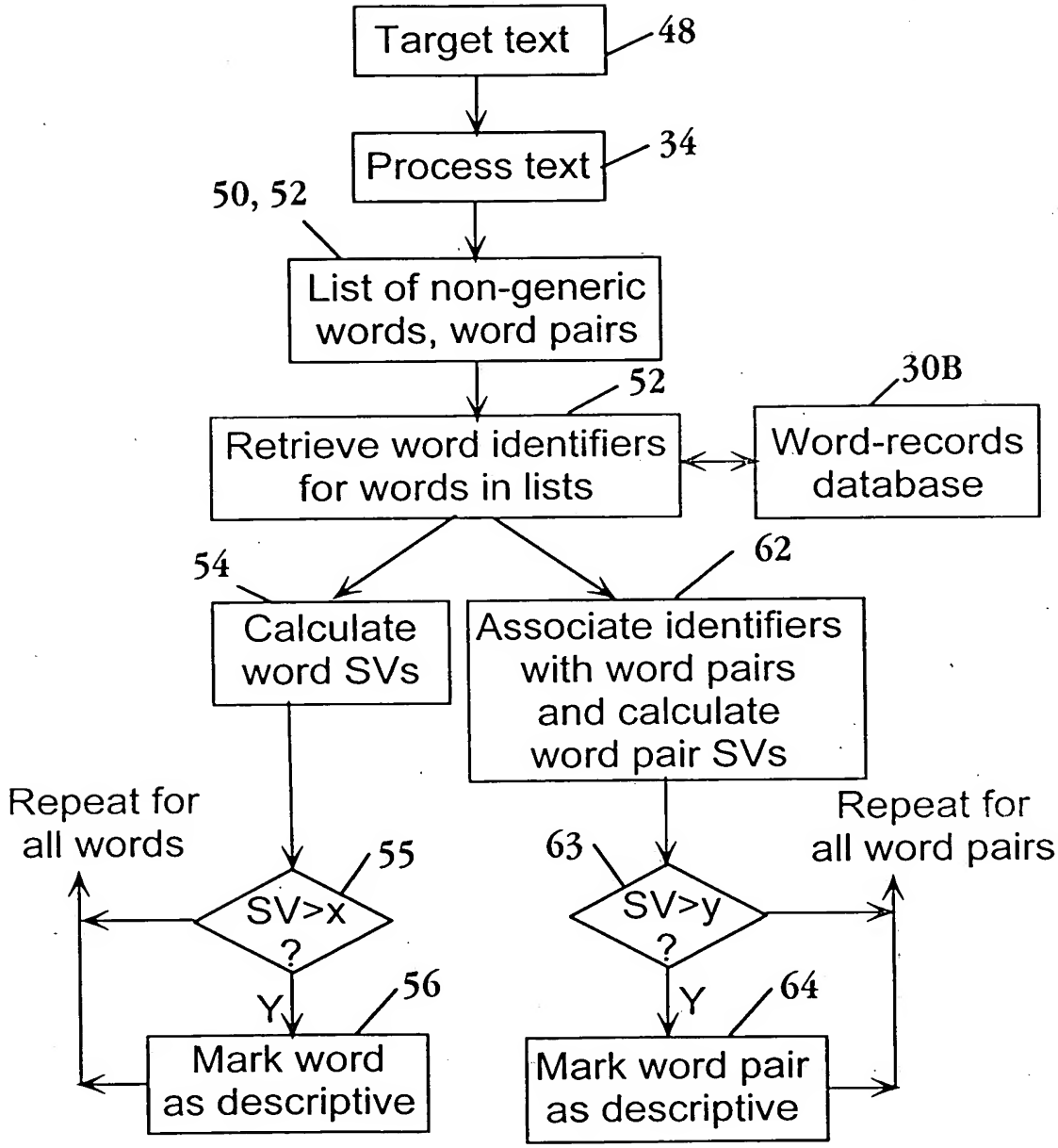

Target text
48
Process text
34
50, 52
List of non-generic words, word pairs
52
30B
Retrieve word identifiers for words in lists
Word-records database
54
62
Calculate word SVs
Associate identifiers with word pairs and calculate word pair SVs
Repeat for all words
55
63
Repeat for all word pairs
SV>x ?
SV>y ?
Y
56
Y
64
Mark word as descriptive
Mark word pair as descriptive

Fig. 3

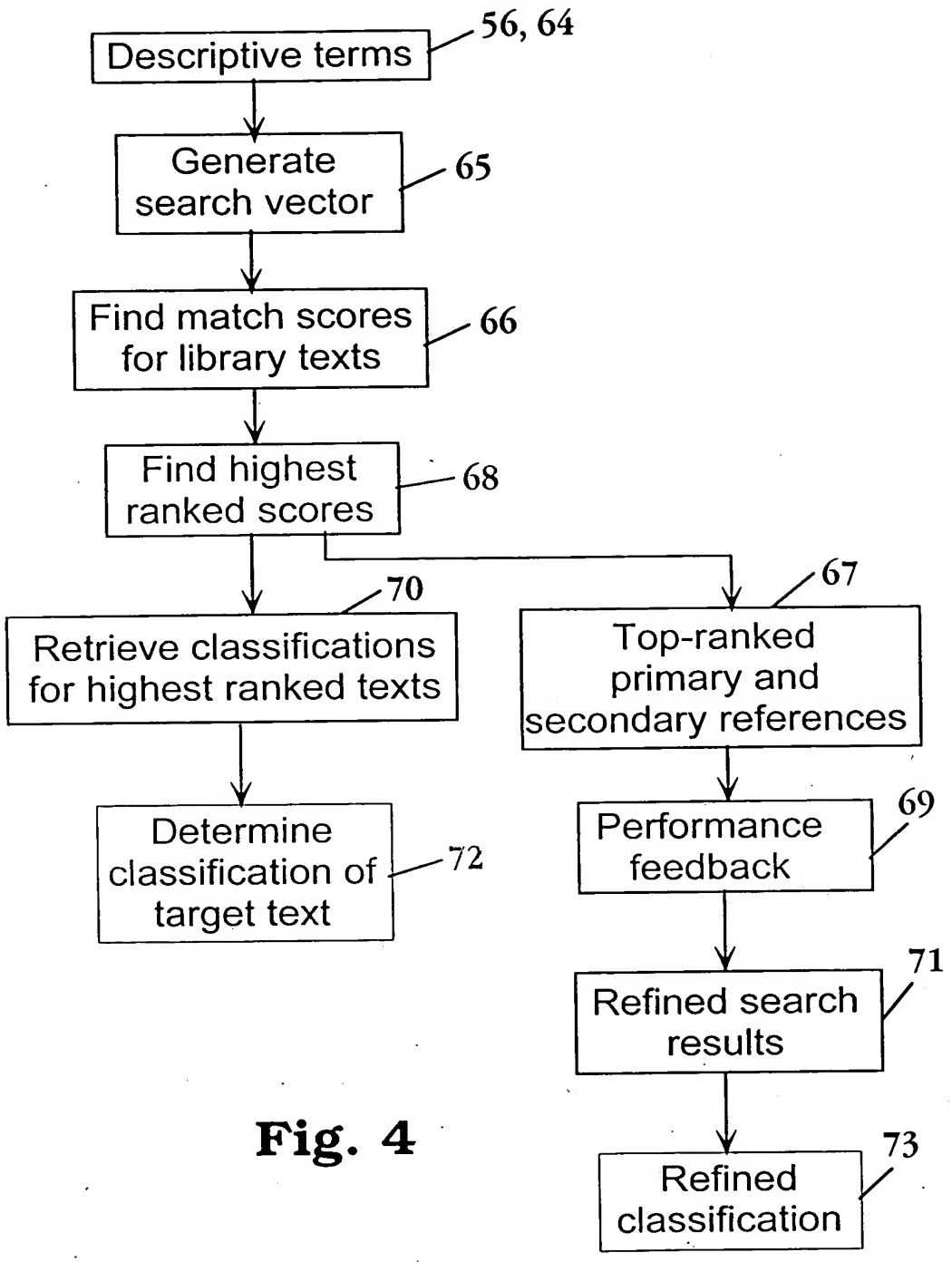

Descriptive terms
56, 64
Generate search vector
65
Find match scores for library texts
66
Find highest ranked scores
68
Retrieve classifications for highest ranked texts
70
Top-ranked primary and secondary references
67
Determine classification of target text
72
Performance feedback
69
Refined search results
71
Refined classification
73

Fig. 4

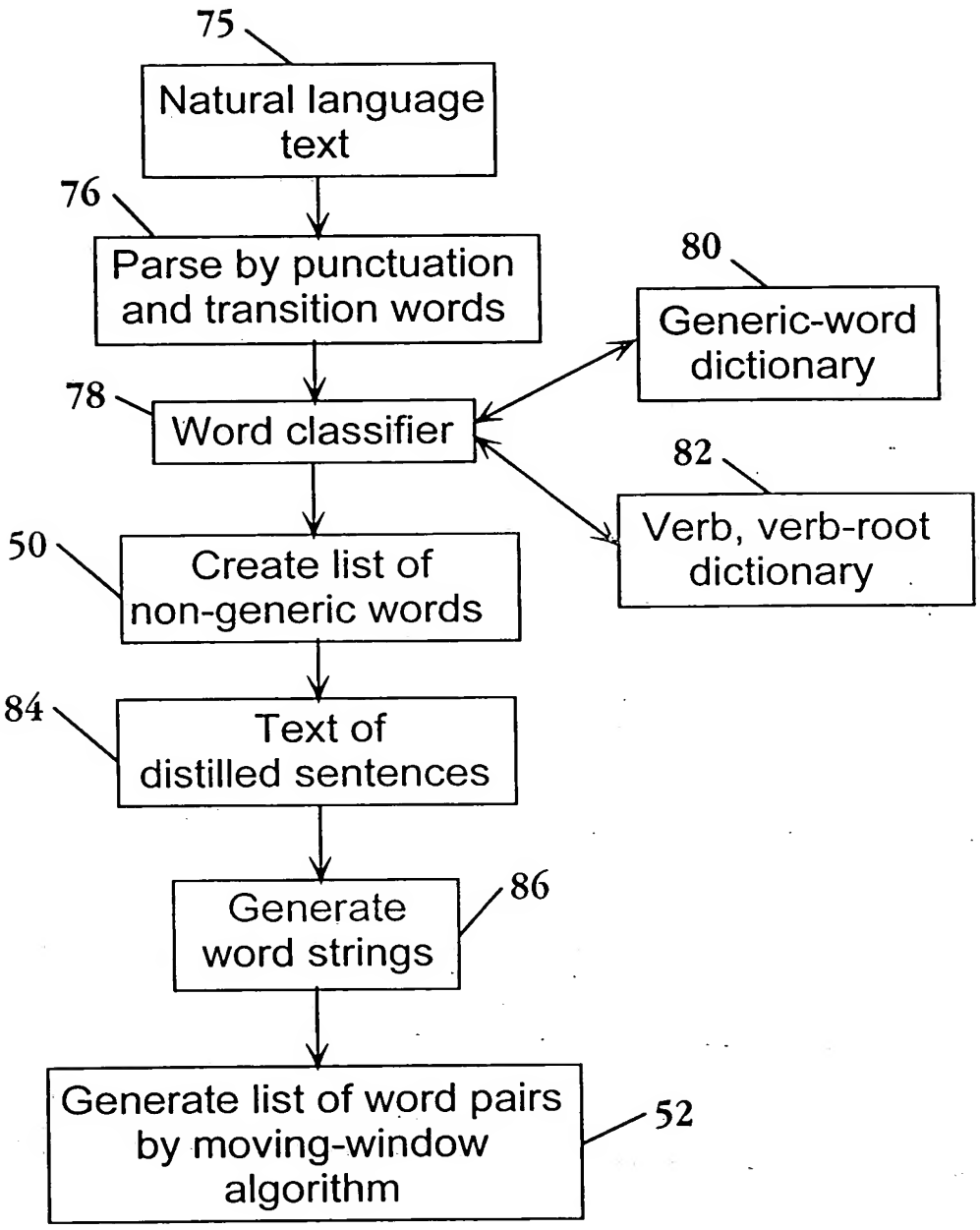

75
Natural language text
76
Parse by punctuation and transition words
80
Generic-word dictionary
78
Word classifier
82
Verb, verb-root dictionary
50
Create list of non-generic words
84
Text of distilled sentences
86
Generate word strings
Generate list of word pairs by moving-window algorithm
52

Fig. 5

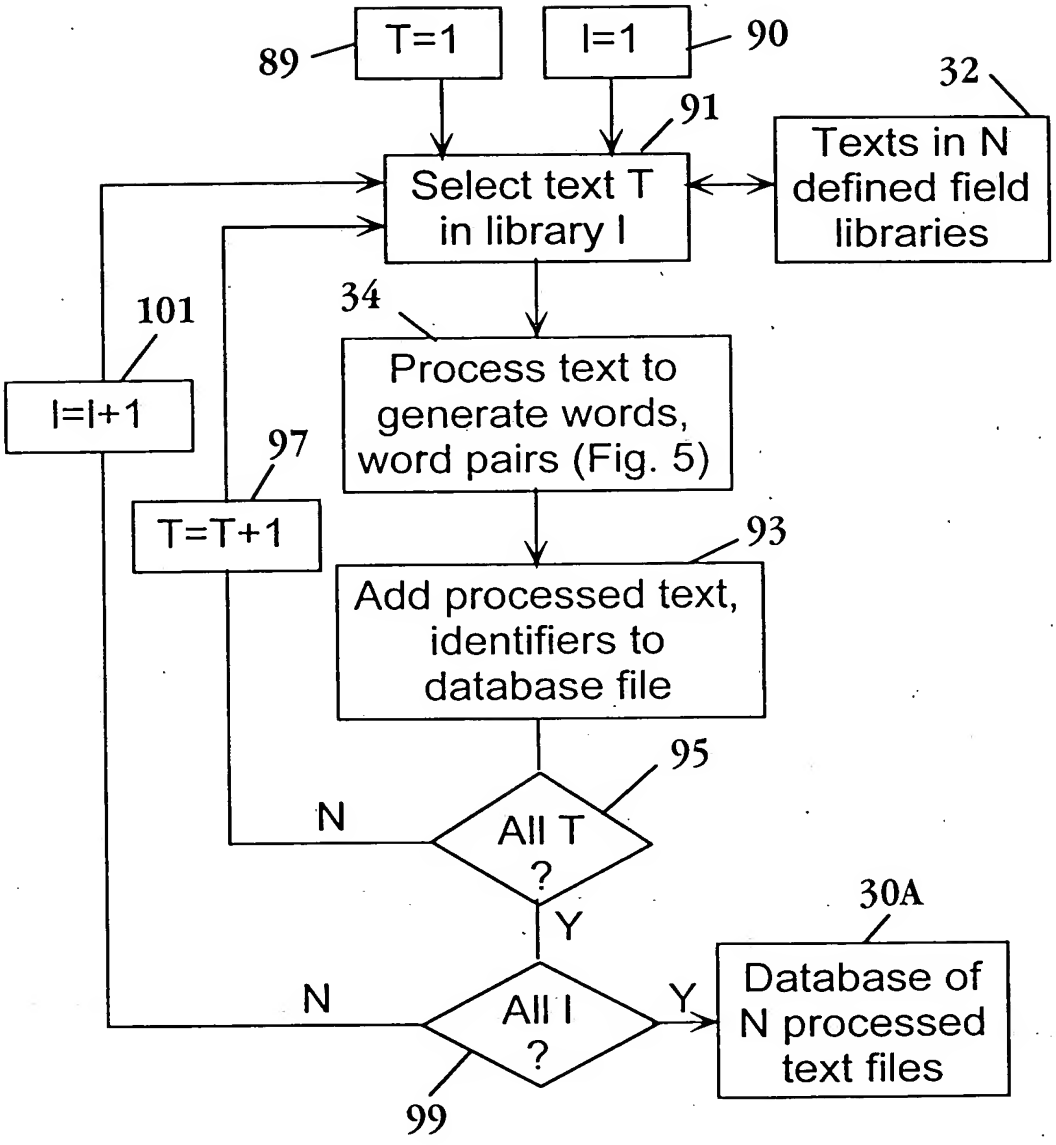

T=1
I=1
89
90
91
32
Texts in N defined field libraries
Select text T in library I
34
Process text to generate words, word pairs (Fig. 5)
101
I=I+1
97
T=T+1
93
Add processed text, identifiers to database file
95
N
All T ?
Y
30A
Database of N processed text files
N
All I ?
Y
99

Fig. 6

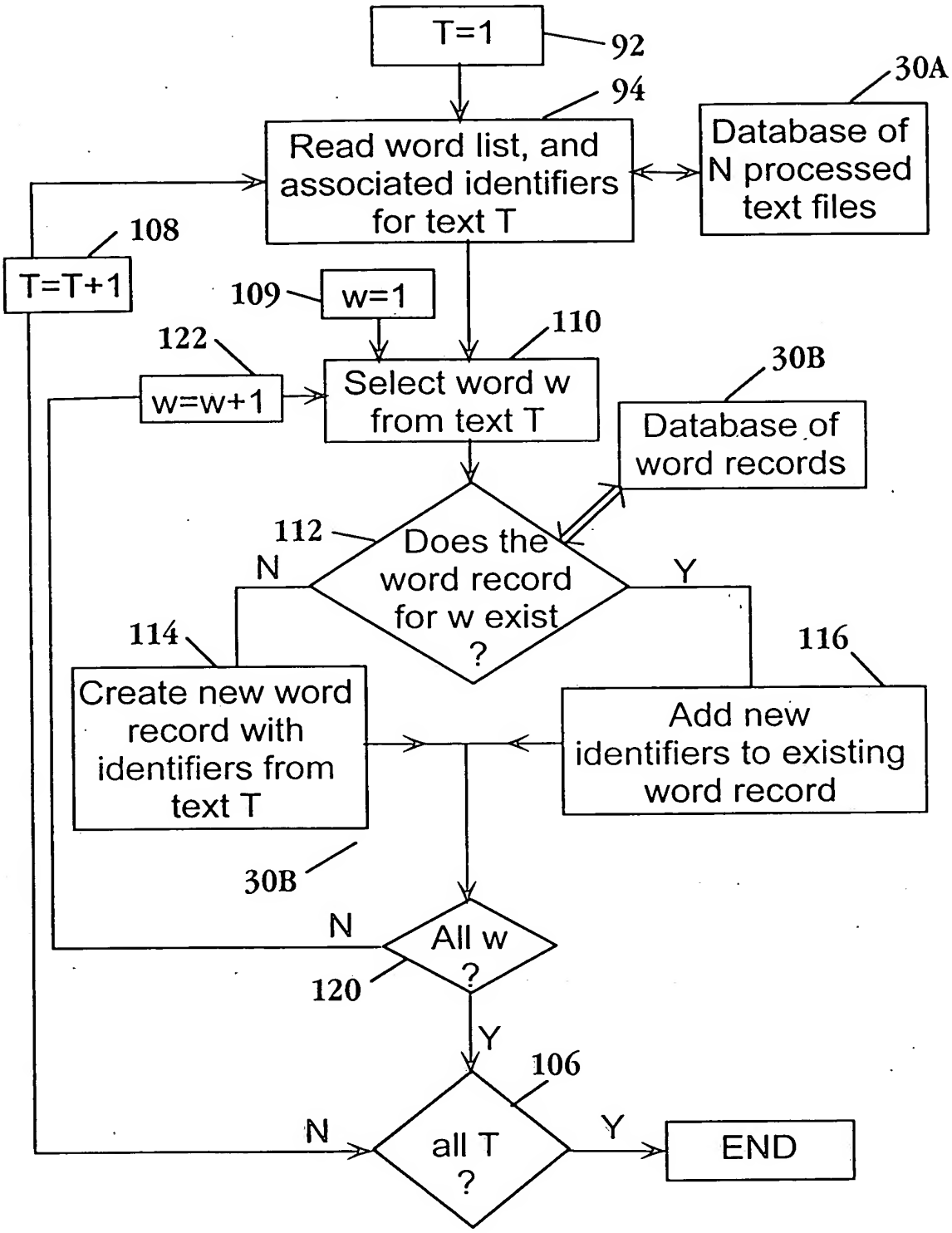

T=1
92
30A
94
Read word list, and associated identifiers for text T
Database of N processed text files
108
T=T+1
109
w=1
110
122
w=w+1
Select word w from text T
30B
Database of word records
112
Does the word record for w exist ?
N
Y
114
116
Create new word record with identifiers from text T
Add new identifiers to existing word record
30B
N
All w ?
120
Y
106
N
all T ?
Y
END

Fig. 7

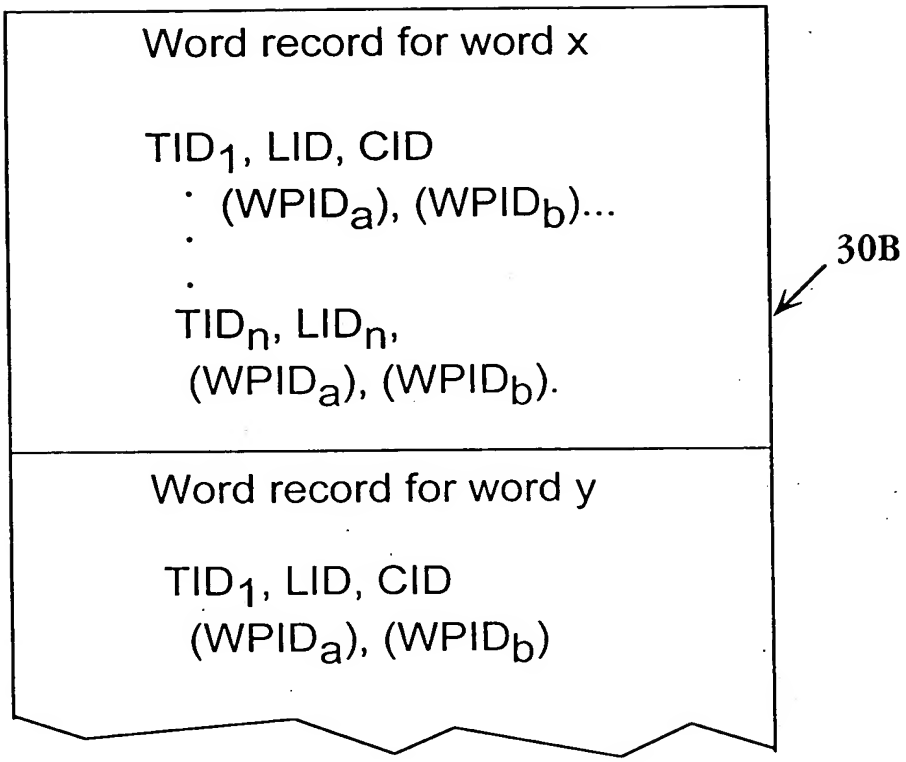

Word record for word x
TID_1, LID, CID
$(WPID_a)$, $(WPID_b)$...
TID_n, LID_n,
$(WPID_a)$, $(WPID_b)$.
30B
Word record for word y
TID_1, LID, CID
$(WPID_a)$, $(WPID_b)$

Fig. 8

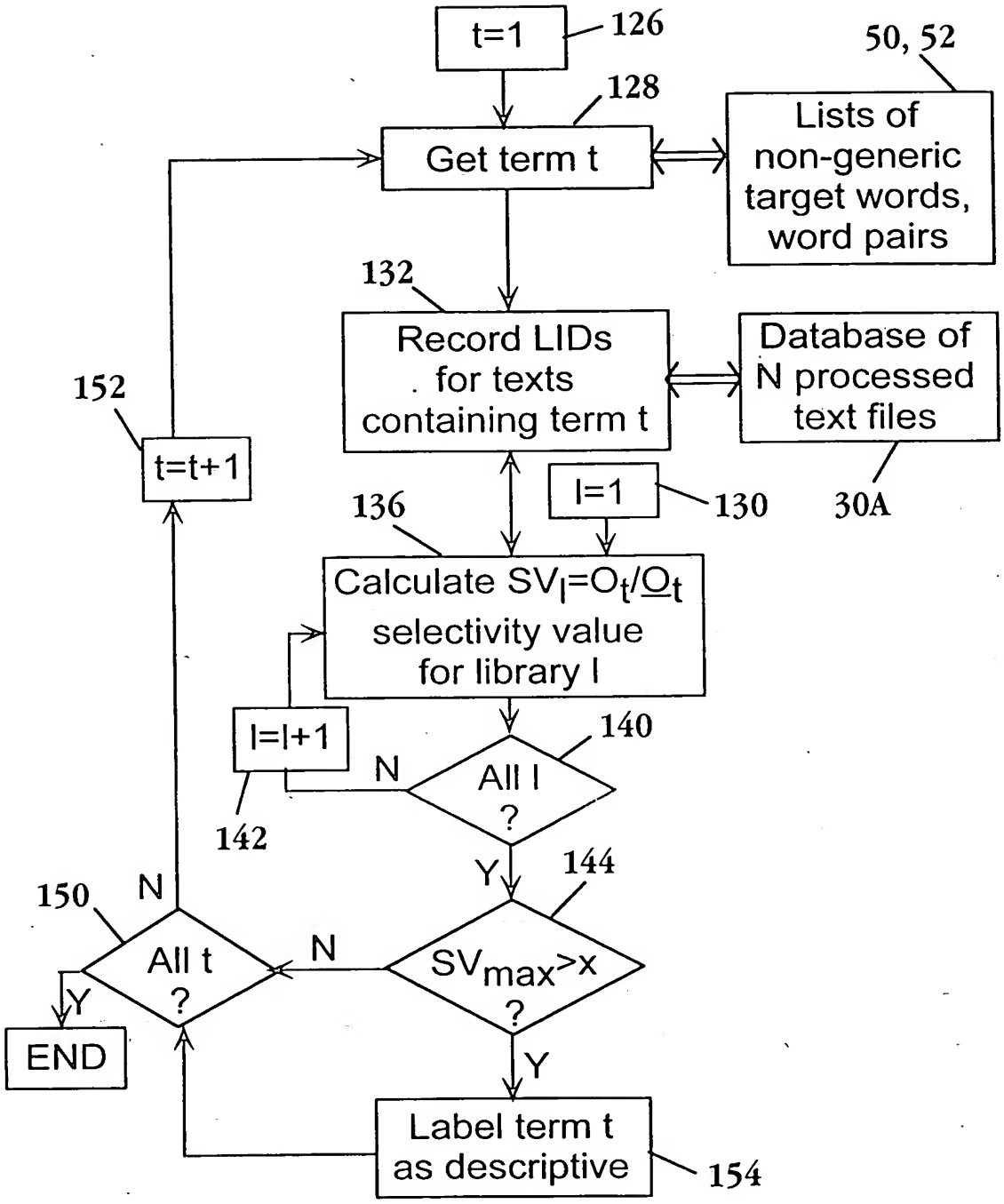

t=1
126
50, 52
128
Get term t
Lists of non-generic target words, word pairs
132
Record LIDs for texts containing term t
Database of N processed text files
152
t=t+1
l=1
130
30A
136
Calculate $SV_l=O_t/\underline{O}_t$ selectivity value for library l
l=l+1
140
N
All l ?
142
Y
144
150
N
All t ?
N
$SV_{max}>x$?
Y
END
Y
Label term t as descriptive
154

Fig. 9A

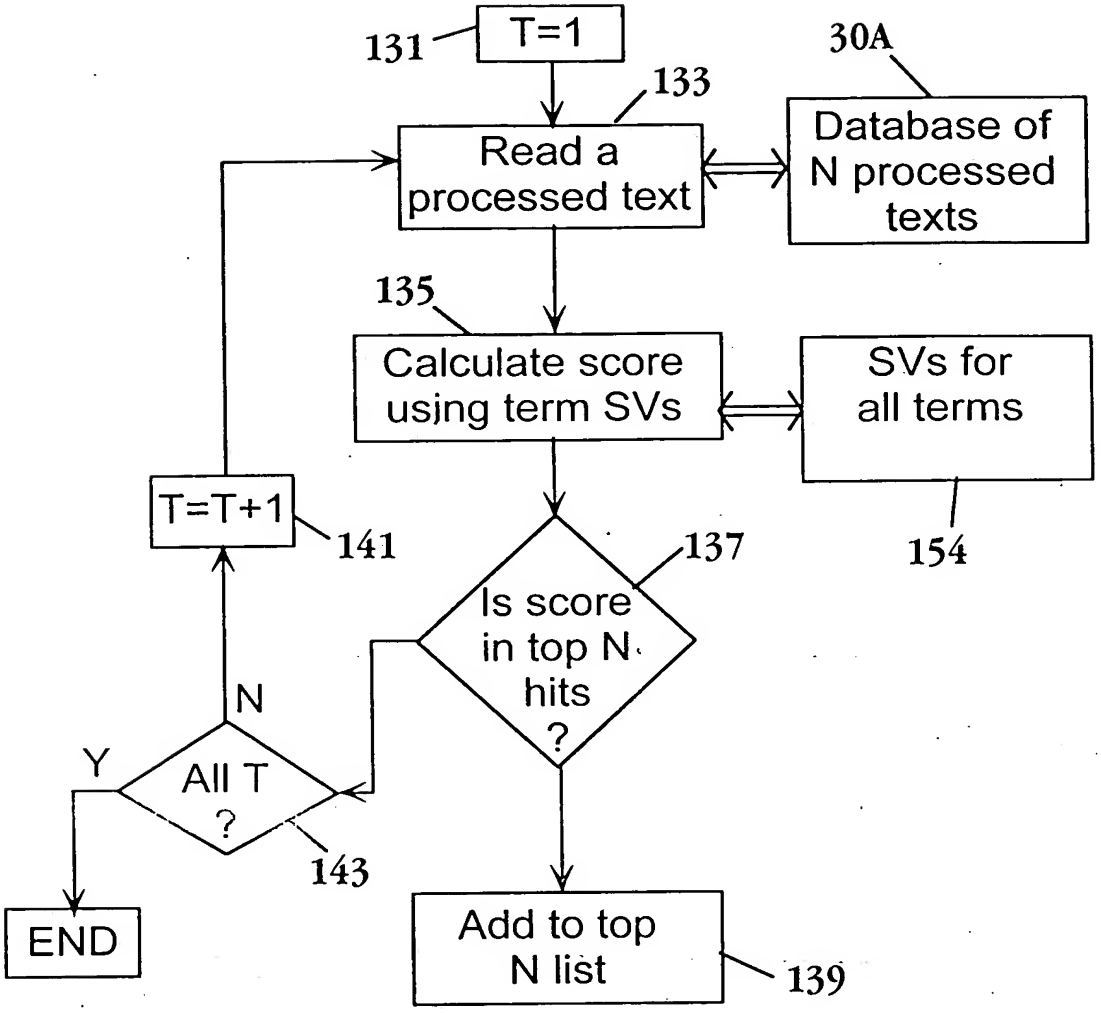

131
T=1
133
Read a processed text
30A
Database of N processed texts
135
Calculate score using term SVs
SVs for all terms
154
T=T+1
141
137
Is score in top N hits ?
N
Y
All T ?
143
END
Add to top N list
139

Fig. 9B

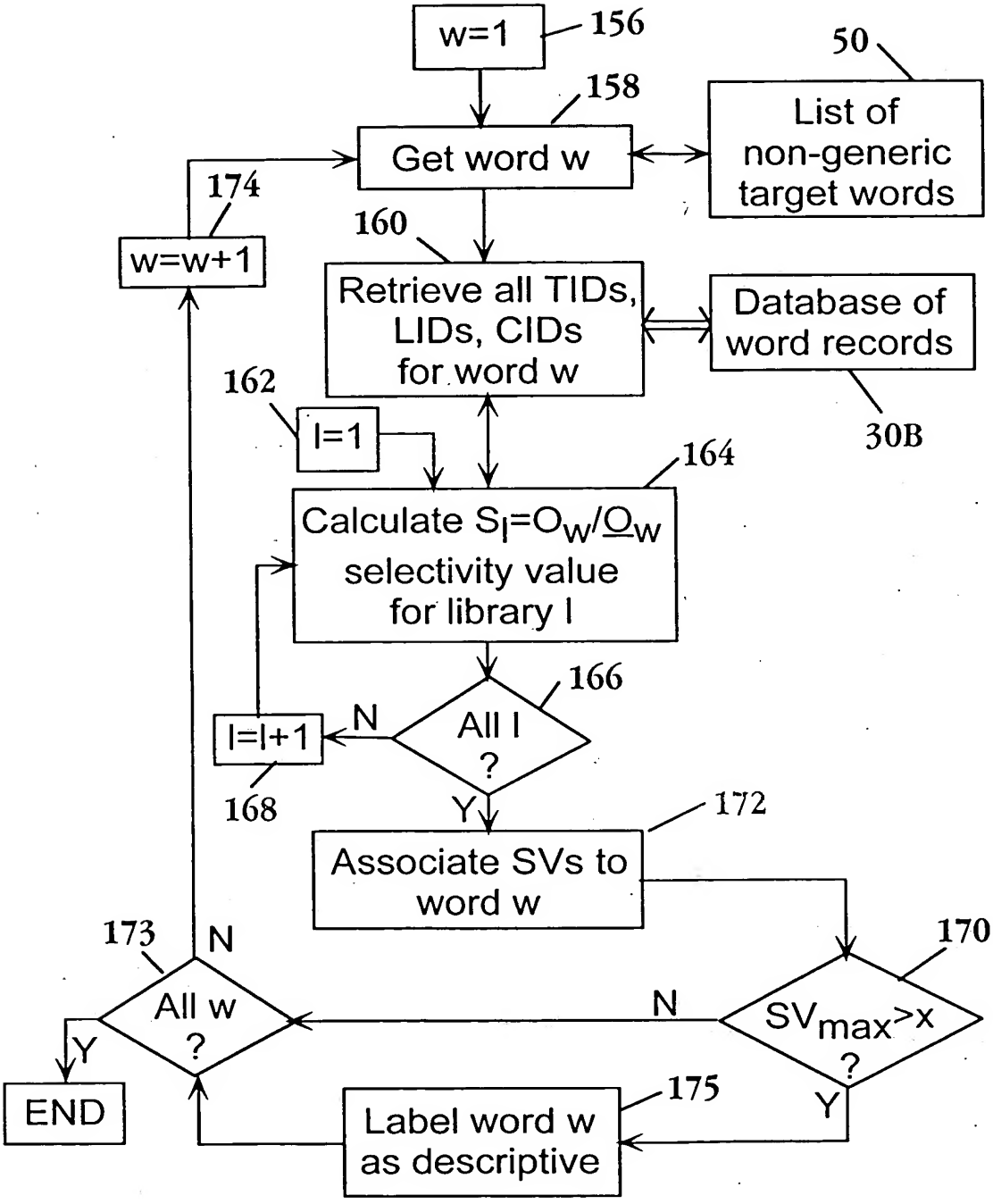

w=1
156
50
158
List of non-generic target words
Get word w
174
160
w=w+1
Retrieve all TIDs, LIDs, CIDs for word w
Database of word records
162
30B
l=1
164
Calculate $S_l=O_w/\underline{O}_w$ selectivity value for library l
166
N
All l ?
l=l+1
168
Y
172
Associate SVs to word w
173
N
170
All w ?
N
$SV_{max}>x$?
Y
END
175
Y
Label word w as descriptive

Fig. 10

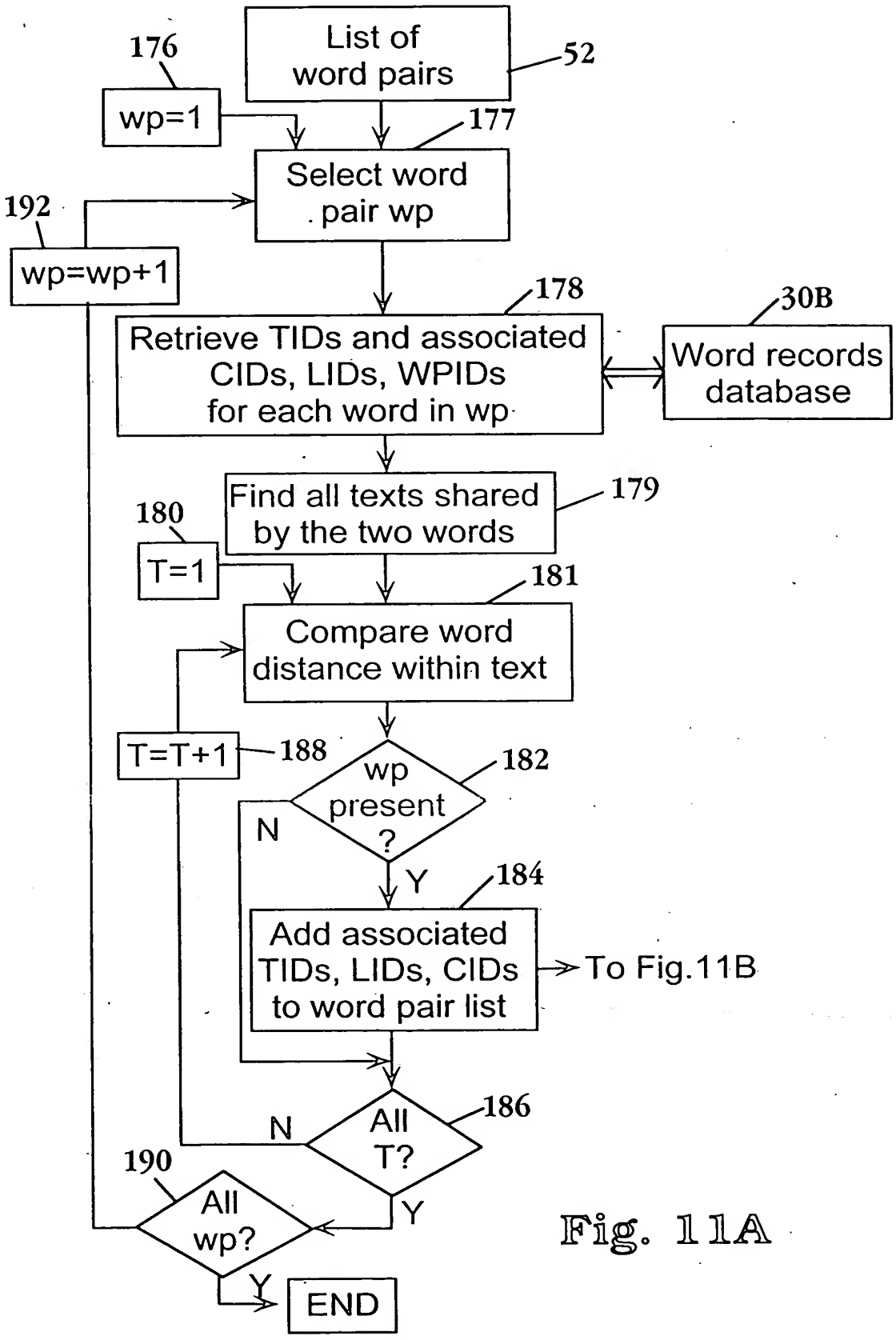

List of word pairs
52
176
wp=1
177
Select word pair wp
192
wp=wp+1
178
Retrieve TIDs and associated CIDs, LIDs, WPIDs for each word in wp
30B
Word records database
Find all texts shared by the two words
179
180
T=1
181
Compare word distance within text
T=T+1
188
wp present ?
182
N
Y
184
Add associated TIDs, LIDs, CIDs to word pair list
To Fig.11B
All T?
186
N
Y
190
All wp?
Y
END

Fig. 11A

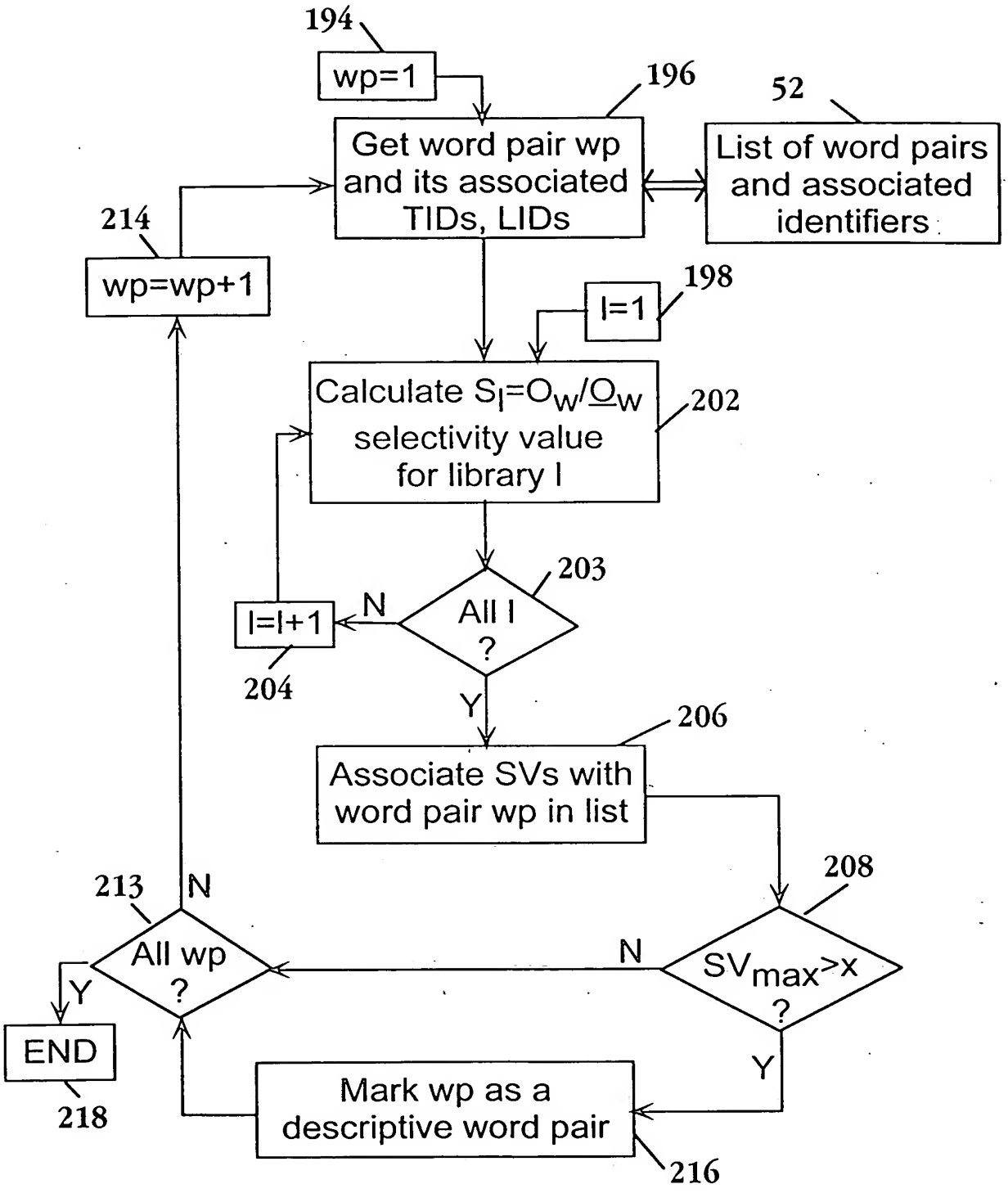

194
wp=1
196
Get word pair wp and its associated TIDs, LIDs
52
List of word pairs and associated identifiers
214
wp=wp+1
198
l=1
Calculate $S_l=O_w/\underline{O}_w$ selectivity value for library l
202
203
All l ?
N
l=l+1
204
Y
206
Associate SVs with word pair wp in list
208
213
N
All wp ?
N
$SV_{max}>x$?
Y
END
218
Y
Mark wp as a descriptive word pair
216

Fig. 11B

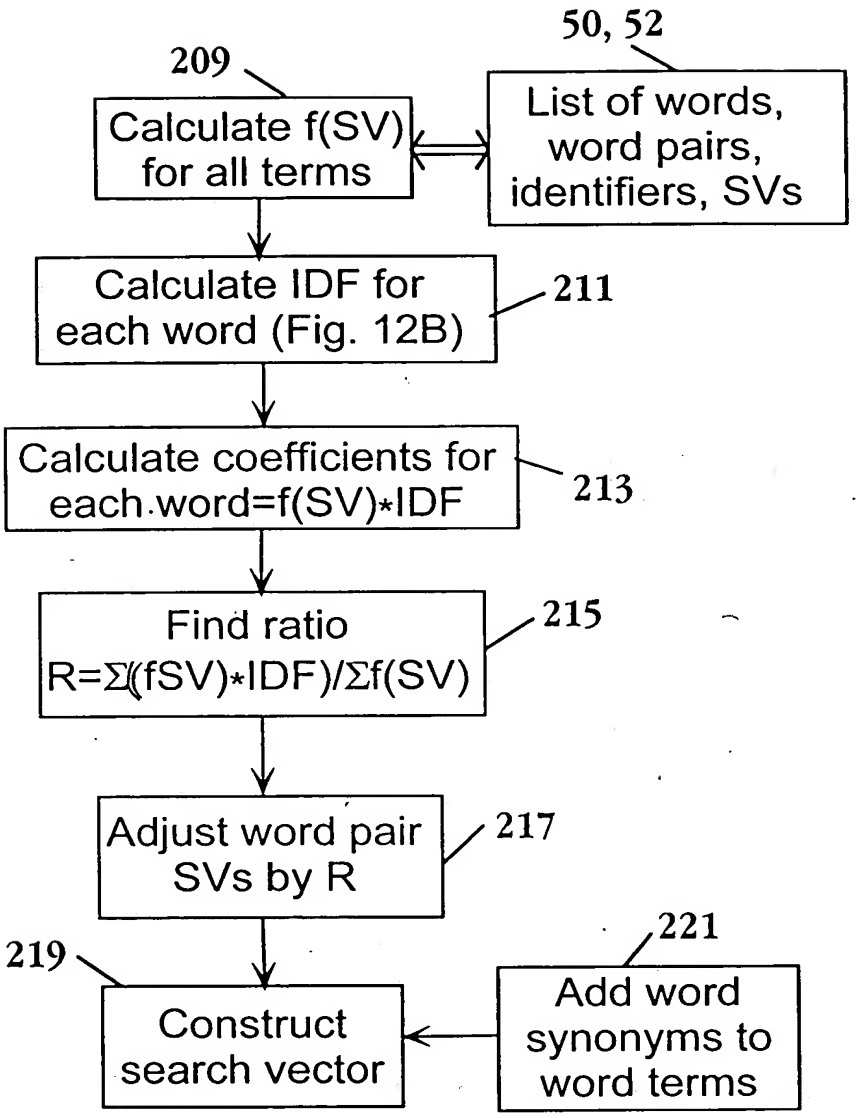

50, 52
209
Calculate f(SV) for all terms
List of words, word pairs, identifiers, SVs
Calculate IDF for each word (Fig. 12B)
211
Calculate coefficients for each word=f(SV)*IDF
213
Find ratio R=Σ(fSV)*IDF)/Σf(SV)
215
Adjust word pair SVs by R
217
219
221
Construct search vector
Add word synonyms to word terms

Fig. 12A

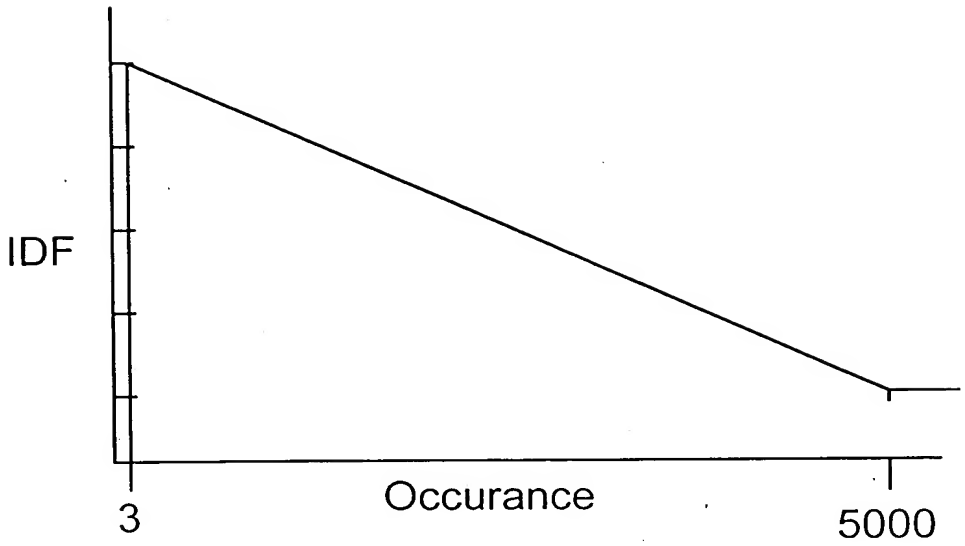


Fig. 12B

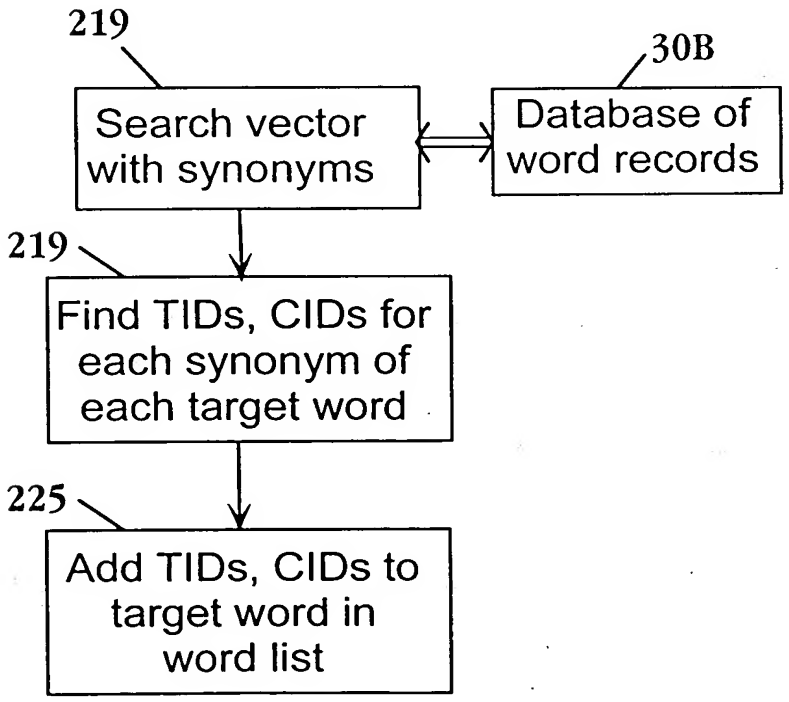


Fig. 12C

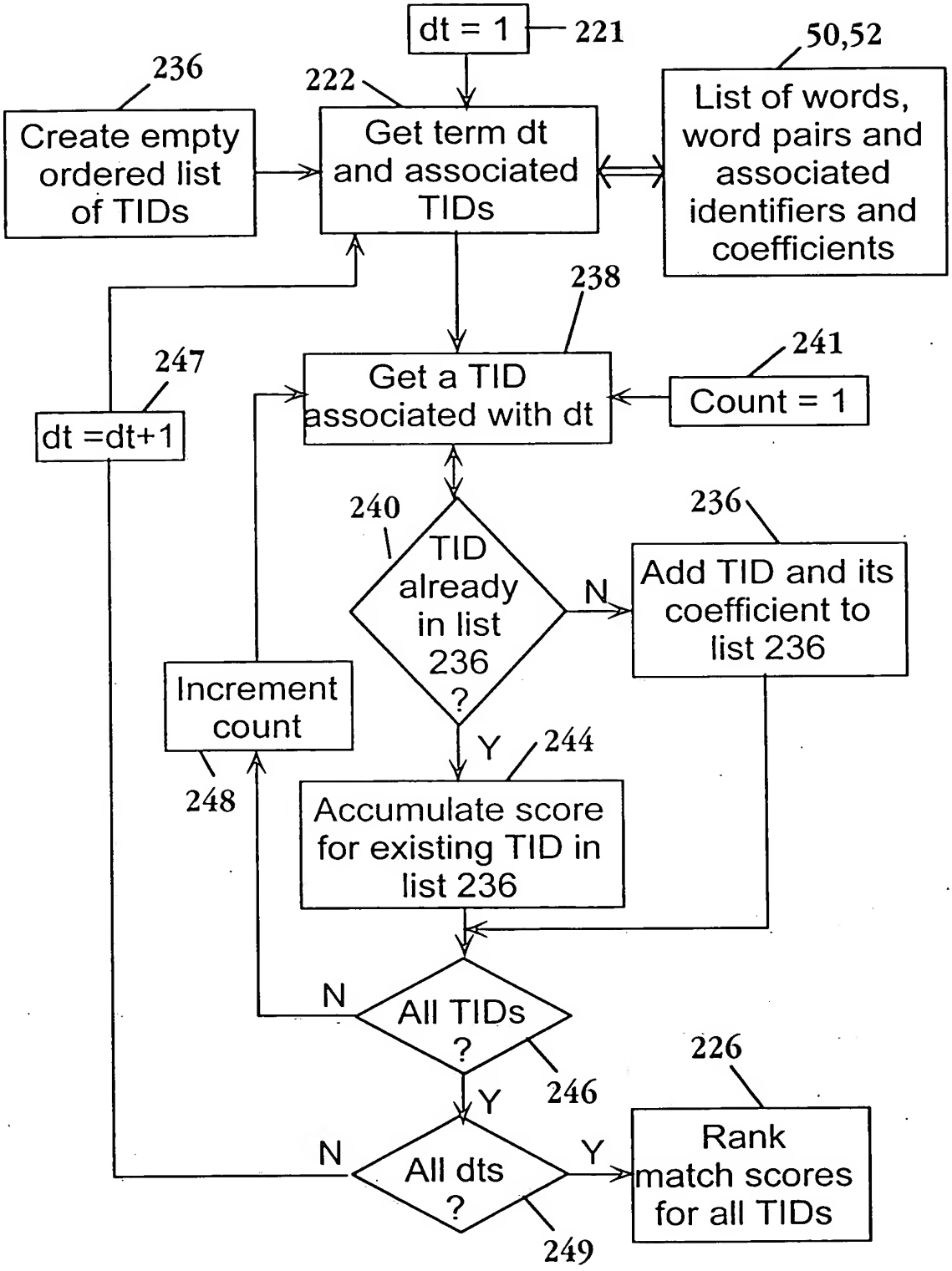

dt = 1
221
50,52
236
222
Create empty ordered list of TIDs
Get term dt and associated TIDs
List of words, word pairs and associated identifiers and coefficients
238
247
241
Get a TID associated with dt
Count = 1
dt =dt+1
240
236
TID already in list 236 ?
N
Add TID and its coefficient to list 236
Increment count
Y
244
248
Accumulate score for existing TID in list 236
N
All TIDs ?
246
226
Y
N
All dts ?
Y
Rank match scores for all TIDs
249

Fig. 13

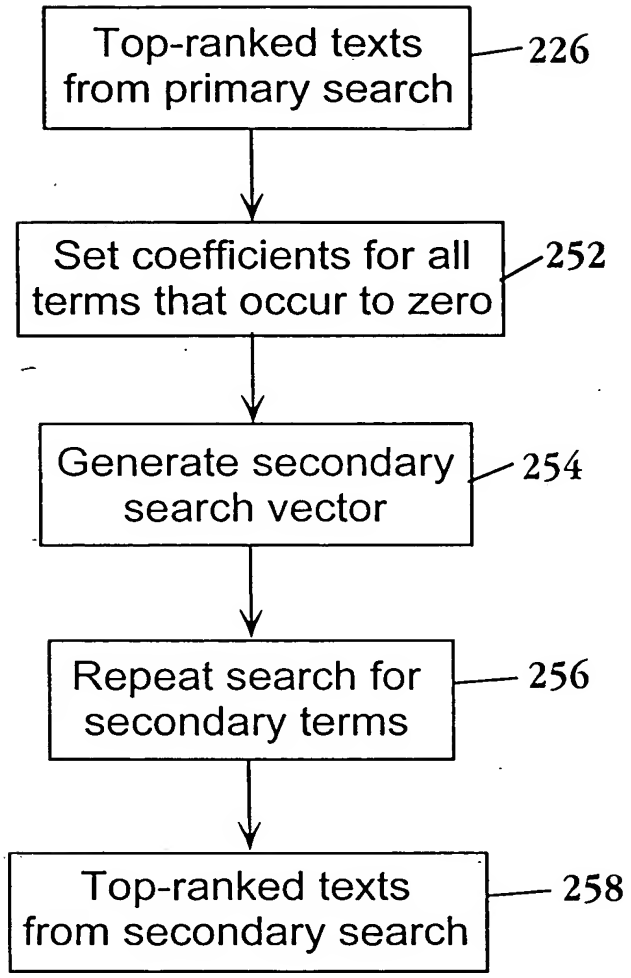

Top-ranked texts from primary search
226
Set coefficients for all terms that occur to zero
252
Generate secondary search vector
254
Repeat search for secondary terms
256
Top-ranked texts from secondary search
258

Fig. 14

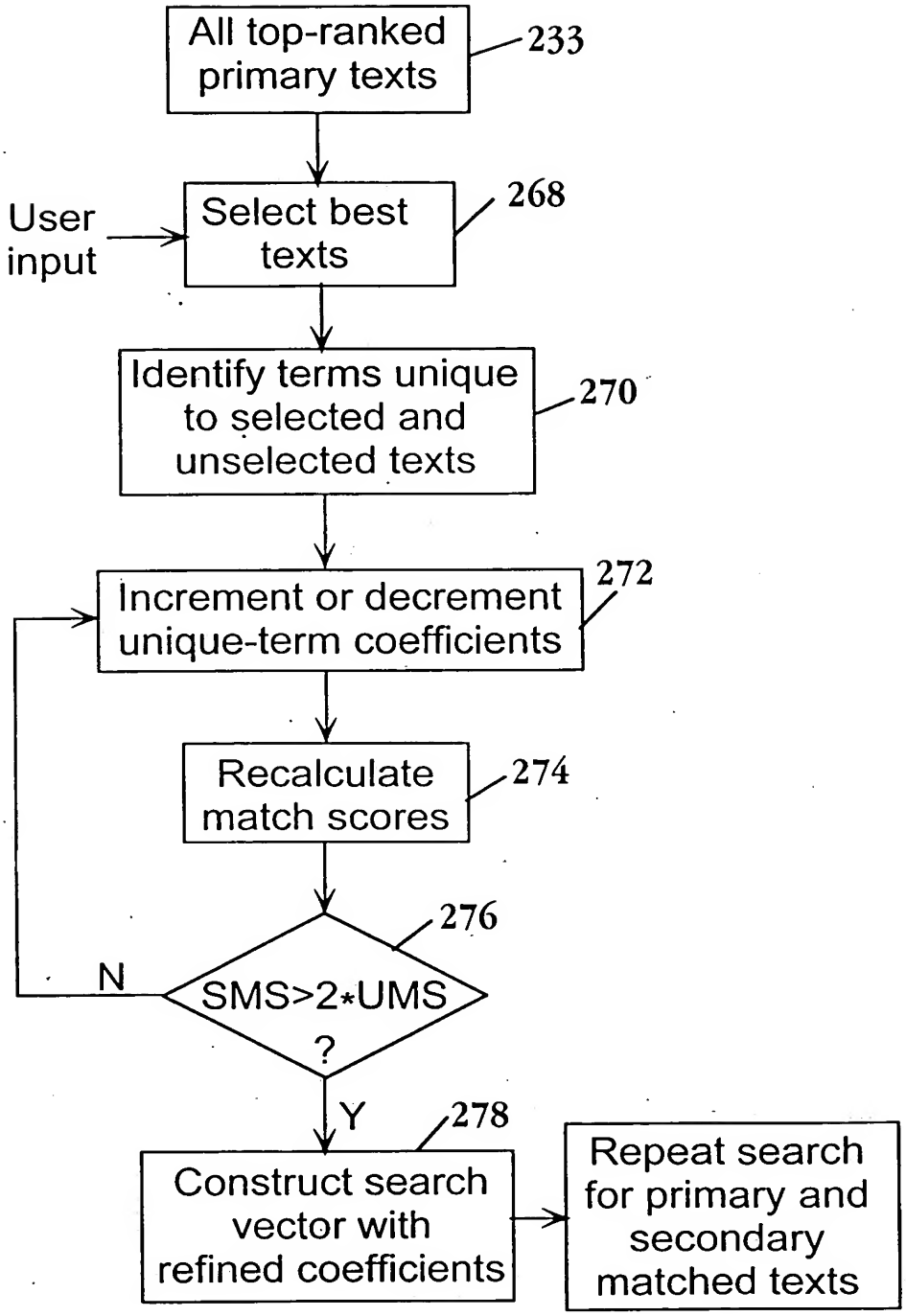

All top-ranked primary texts
233
User input
Select best texts
268
Identify terms unique to selected and unselected texts
270
Increment or decrement unique-term coefficients
272
Recalculate match scores
274
276
SMS>2*UMS ?
N
Y
278
Construct search vector with refined coefficients
Repeat search for primary and secondary matched texts

Fig. 15

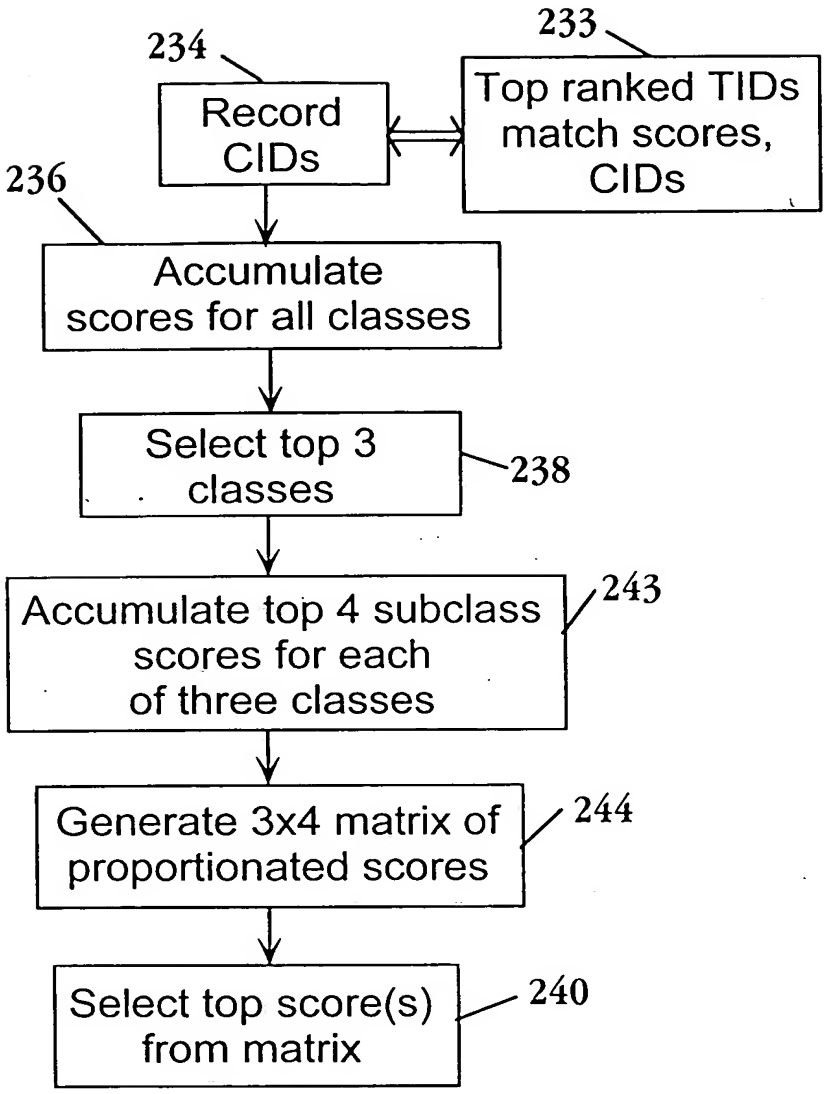
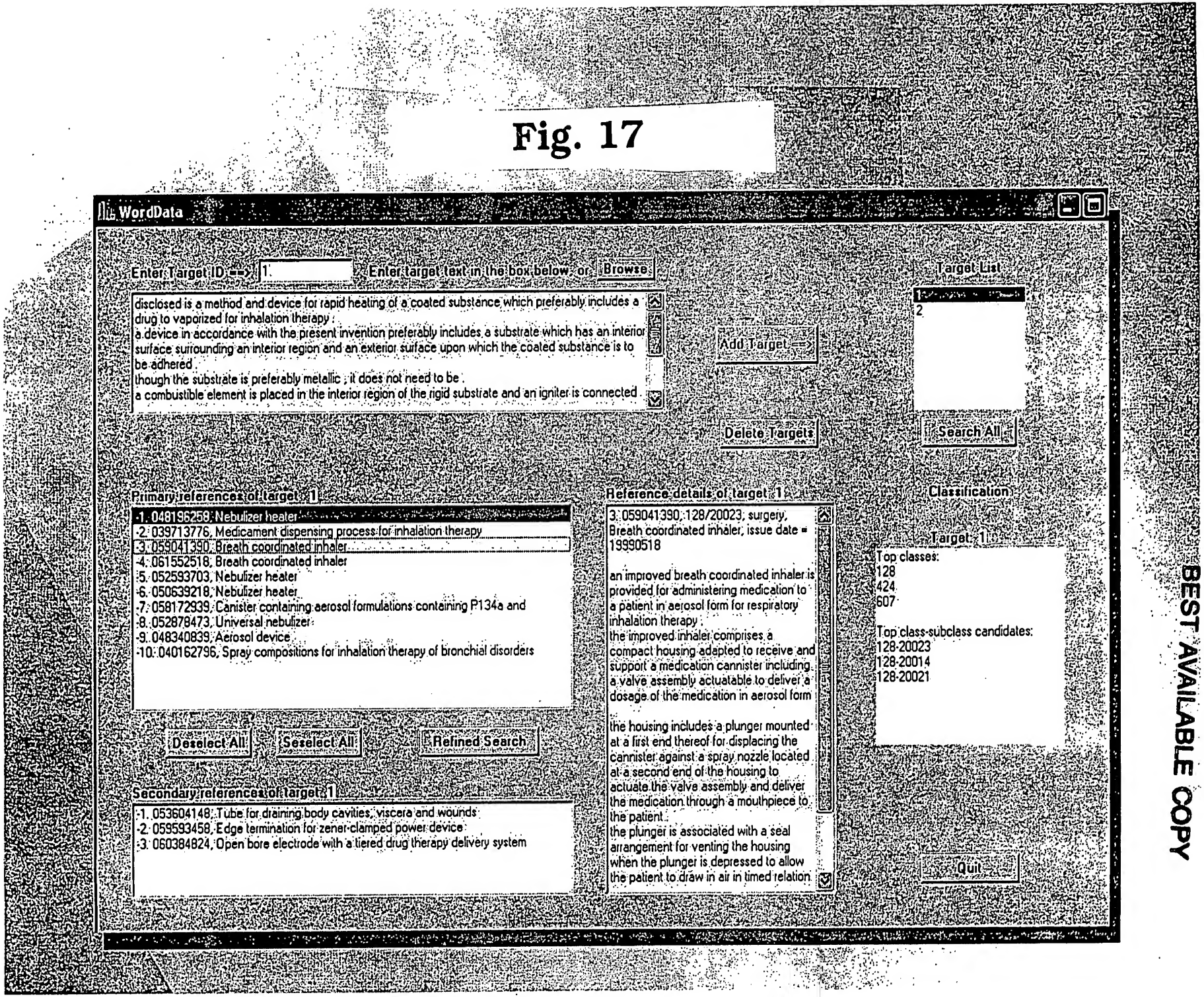

234
233
Record CIDs
Top ranked TIDs match scores, CIDs
236
Accumulate scores for all classes
Select top 3 classes
238
Accumulate top 4 subclass scores for each of three classes
243
Generate 3x4 matrix of proportionated scores
244
Select top score(s) from matrix
240

Fig. 16

Fig. 17



Search Results Comparison for Three Search Methods on Five Targets

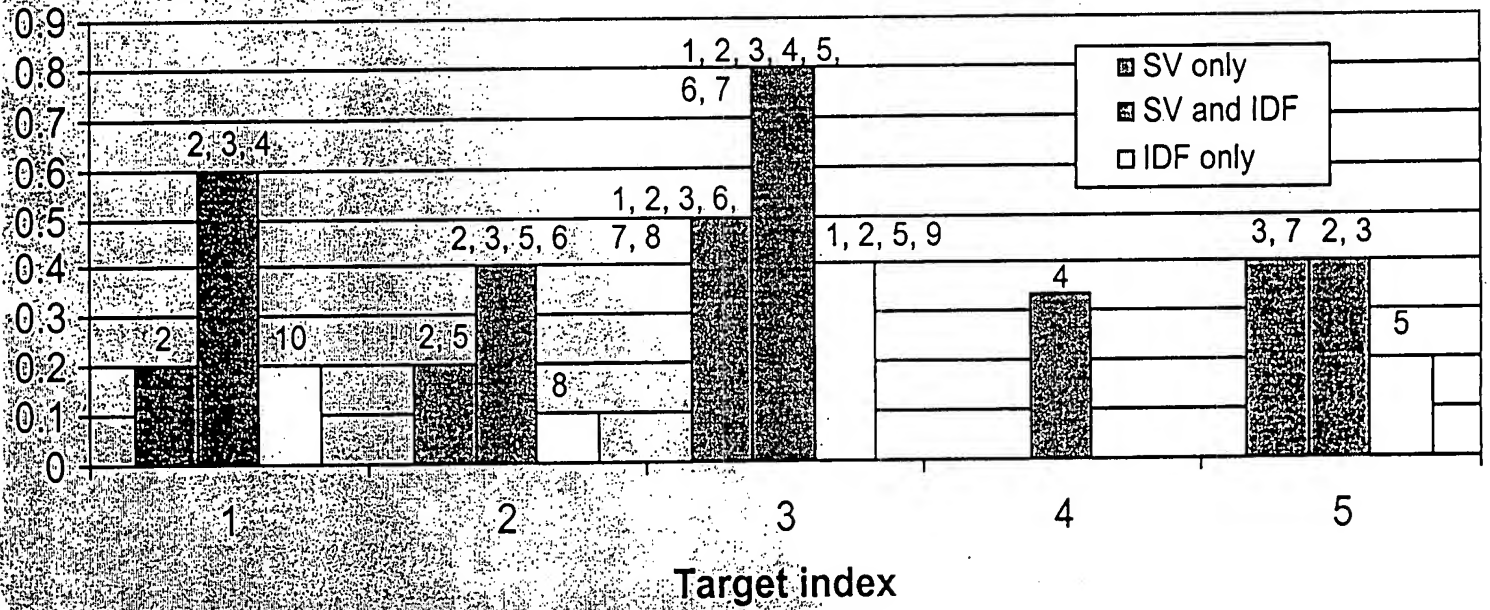


Fig. 18

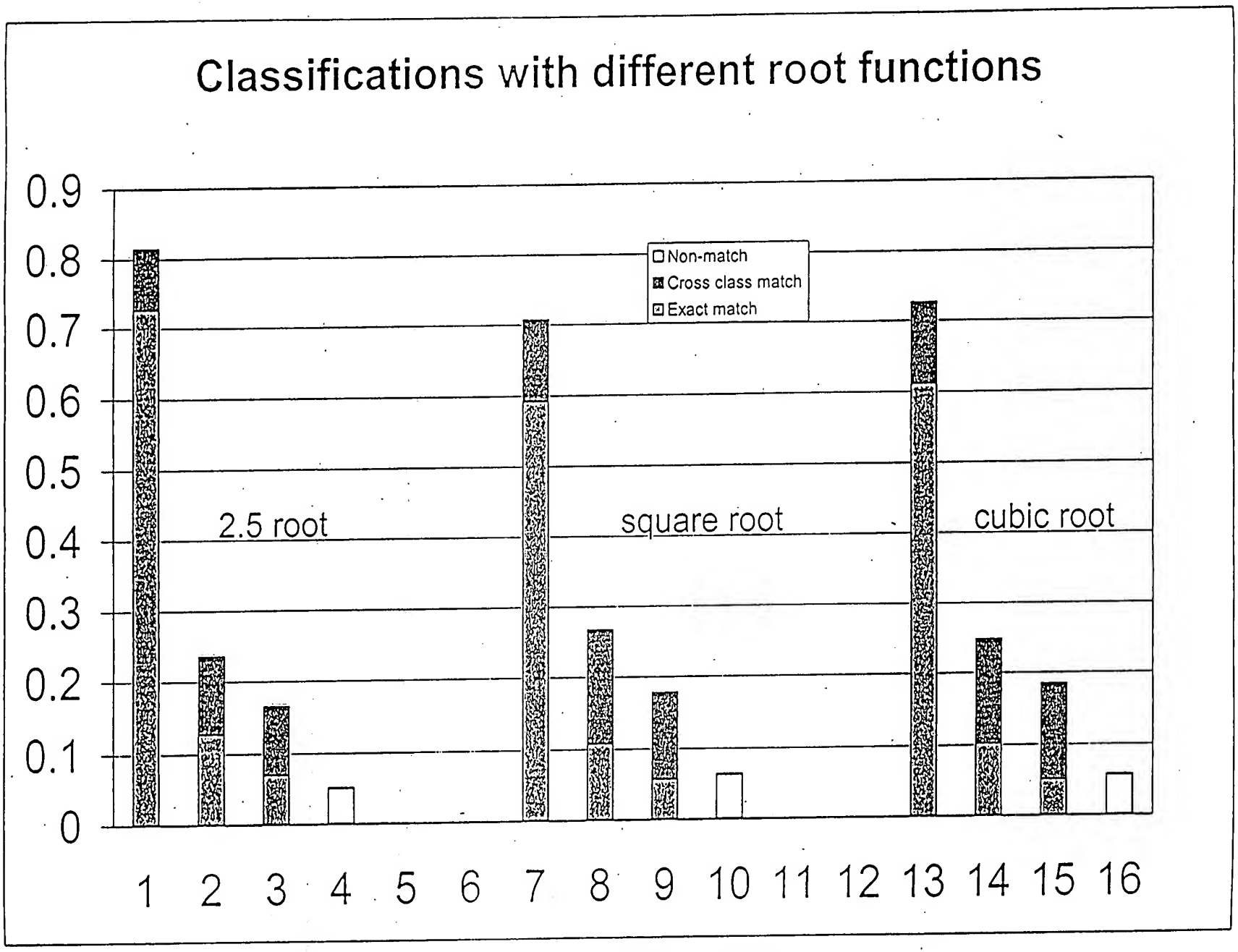

Classifications with different root functions
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Non-match
Cross class match
Exact match
2.5 root
square root
cubic root
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Fig. 19A

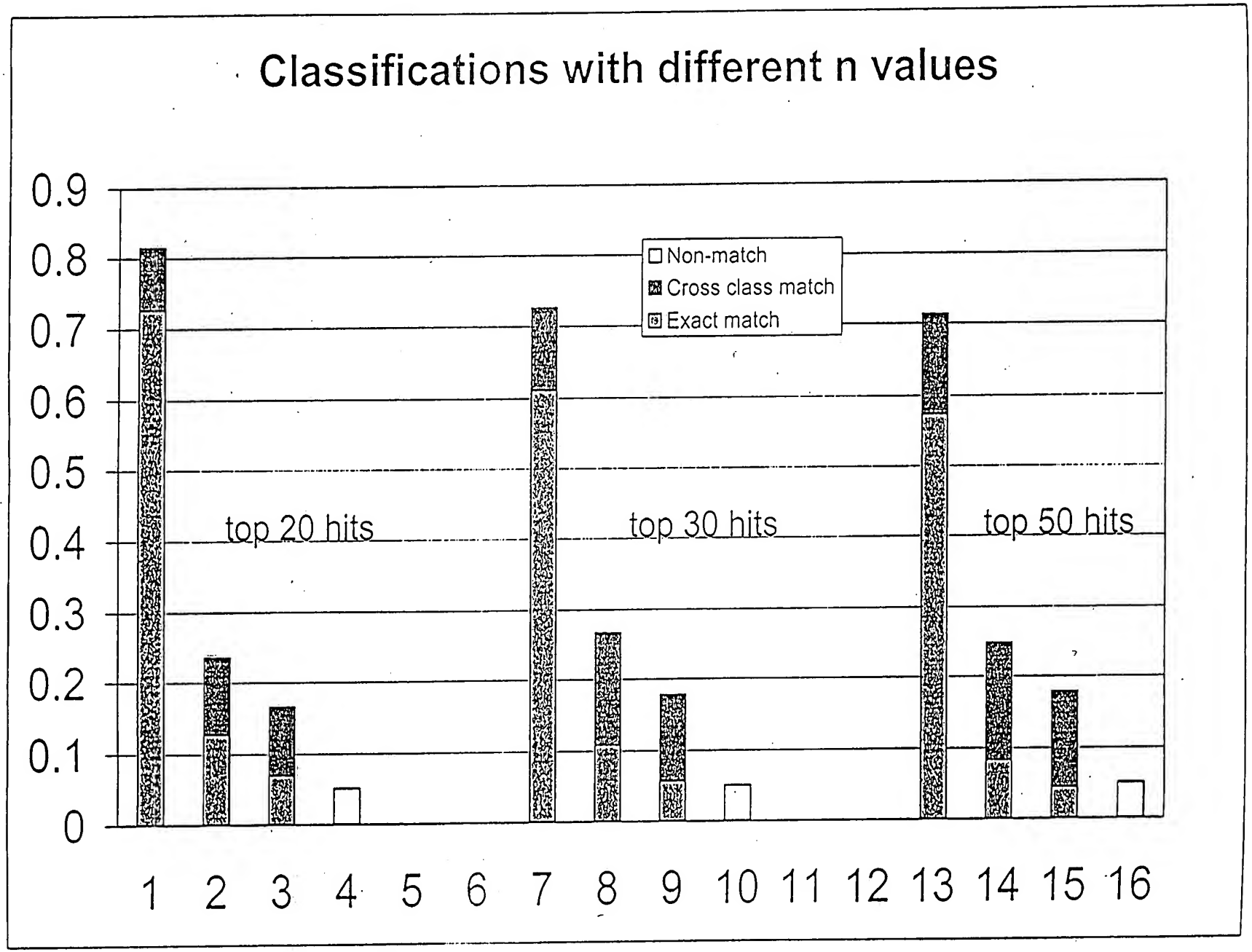

Classifications with different n values
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Non-match
Cross class match
Exact match
top 20 hits
top 30 hits
top 50 hits
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Fig. 19B

Classification Results From Six Approaches

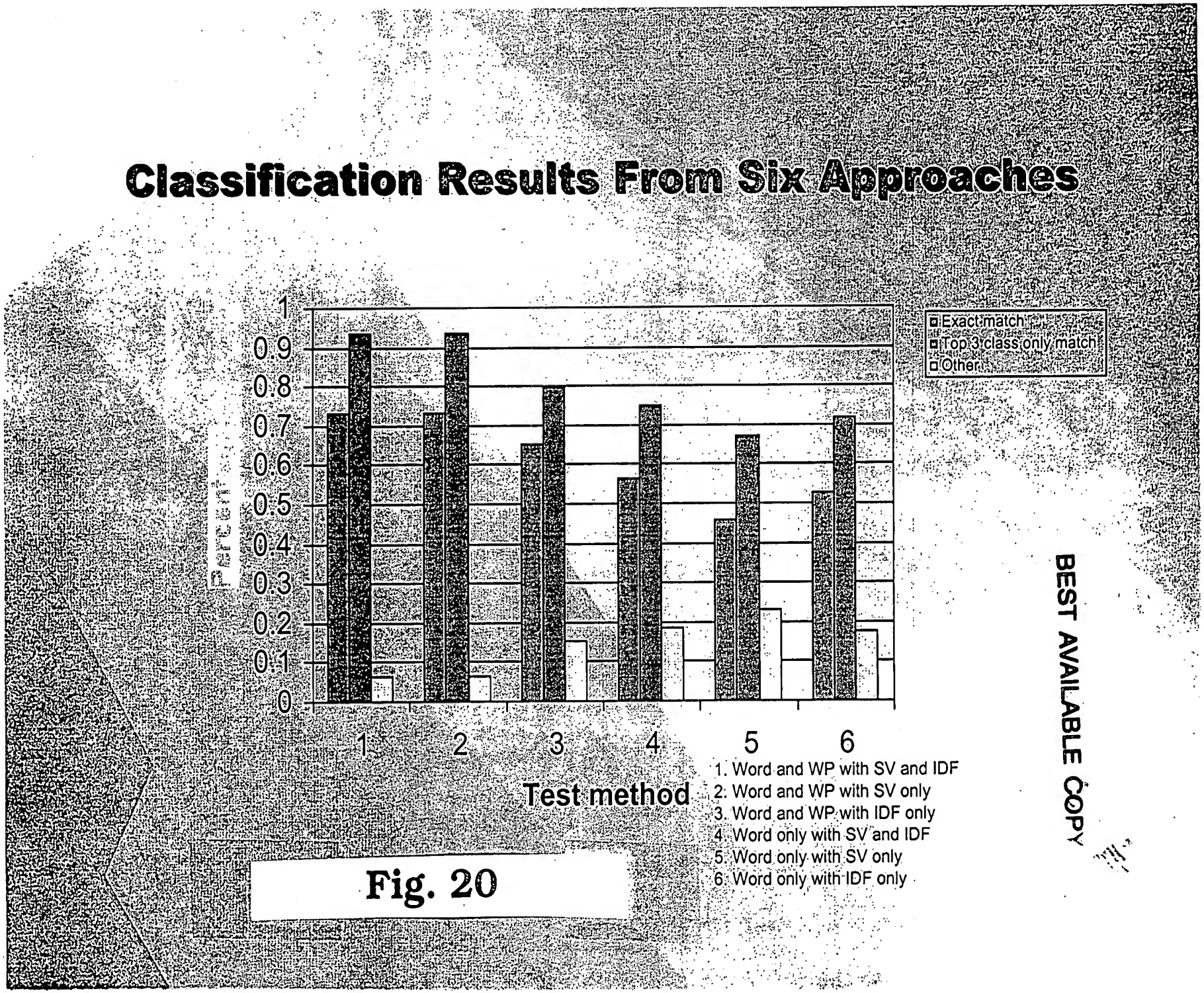
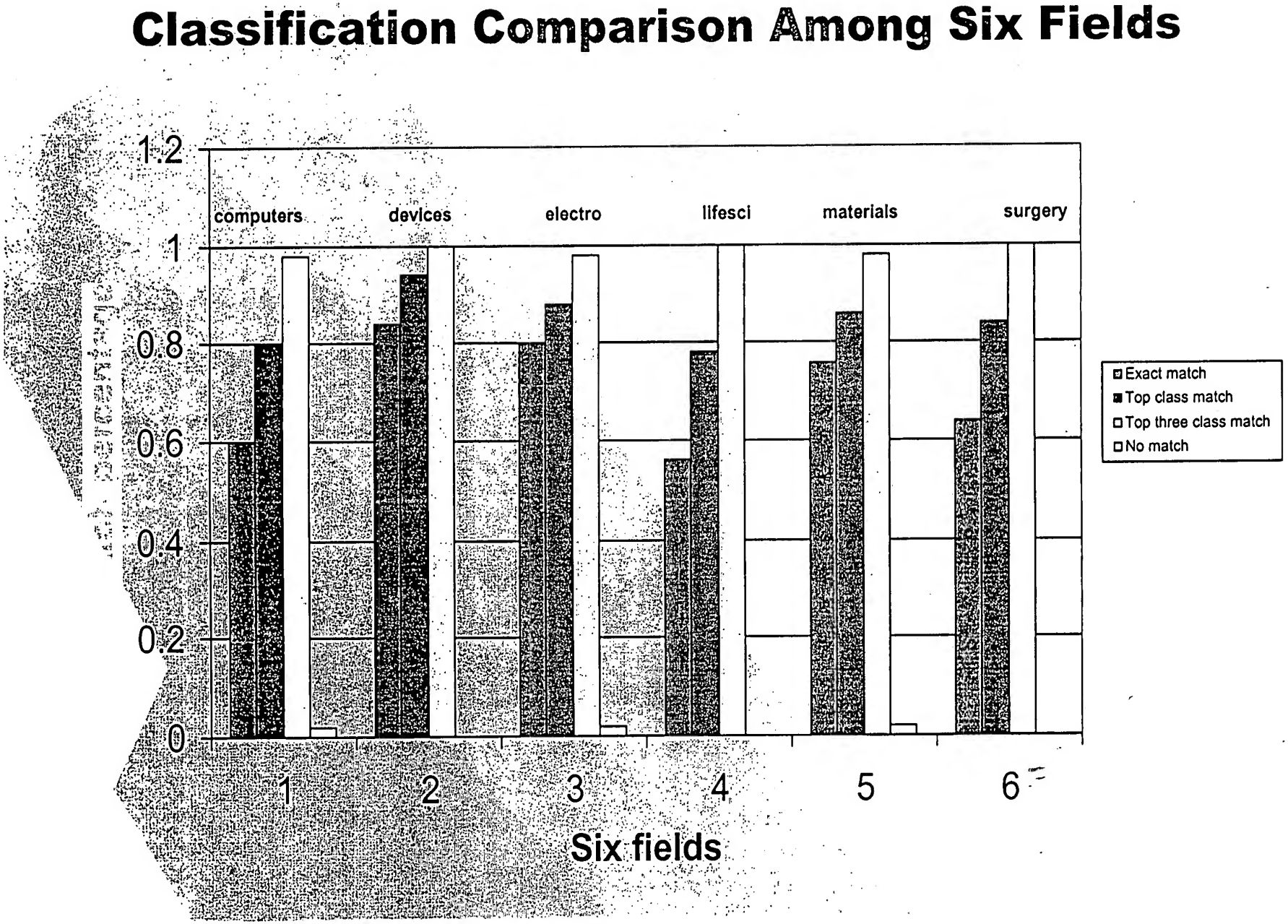


Fig. 20

Fig. 21


Classification Comparison Among Six Fields
1.2
1
0.8
0.6
0.4
0.2
0
computers
devices
electro
lifesci
materials
surgery
1
2
3
4
5
6
Six fields
Exact match
Top class match
Top three class match
No match